Is Singapore's annual fertility rate less affected by marriages?

Ernest Ng Er Rui | B1580997| 5th September 2018

Abstract

To investigate the looming population issues that Singapore faces, the Singapore government identifies the unwillingness to settle down or lack of marriages as a cause for plummeting fertility rates in the country. We aim to verify the cause of low fertility rates in Singapore by examining the association between marriages in Singapore and fertility rates. Using a combination of linear regression and ratio analysis, we derive two measures of correlations to gauge different aspects of the relationship between marriage and fertility in Singapore. The findings reveal that the relationship between marriage and fertility are still functional but not as relevant as before. There are also assumptions used in the approach of this study which needs to be further considered in the findings and other factors that may be more relevant to the fertility rates of Singapore.

1. Datasets

The first dataset used for this research study was 'Marriage Rates, Annual' from Singapore's Department of Statistics. It is a CSV formatted file detailing the marriage rate in Singapore for both genders and the crude marriage rate from 1980 to 2017. The marriage rates in this dataset refers to the percentage of Singaporean male or female married for every 1,000 unmarried Singaporeans in a year.

In this research, we will be using marriage rates for male and females only for the independent variable of marriages. The variable 'crude marriage rates' was not utilised because it was calculated based on resident marriages where both the groom and brides are Singapore residents. Including crude marriages would narrow our definition of marriages in this study as it only reflects a sample rather than the entire Singaporean population. Hence, we will only use male and female marriage rates only to prevent any distortion of our study results.

The second dataset for this research study was 'Births and Fertility Rates, Annual' from Singapore's Department of Statistics. It is also another CSV formatted file which includes data of the annual number of live-births and total fertility rate dated from 1960 to 2017.

In this research, we used the variables 'Total Fertility Rate' and 'Net Reproduction Rate' to reflect the dependent variable 'annual fertility rate'. Total Fertility Rate is the number of live births per female in their reproductive years.Net Reproduction Rate refers to the average number of daughters born per female in their reproductive years.

Both datasets were read into data frames using pandas function, pandas.read_csv() and cleaned accordingly. Both datasets also have years from 1980 to 2017 marked as columns and the years can be common feature to merge both datasets. The treatment of the datasets will be further elaborated in the data cleaning segment of this report.

Both datasets were sourced from https://www.singstat.gov.sg/

2. Research Question

In 2017, the proportion of Singaporeans at age 65 or older was equivalent to the proportion of residents younger than 15 for the first time in modern Singapore history. Demographic issues such as an ageing workforce had begun to loom over the young nation, with their fertility rates recording a low of 0.83; the lowest in the world (Kotecki, 2018).

With an economy that thrives on a talented workforce and foreign investments, the government is in a race against time to reverse this demographic trend for it brings economic repercussions. While the nation's economic outlook hangs in the balance of Singapore's response to an ageing population, the government has often mooted reasons for the low fertility rates to be Singaporeans' "lack of dating" and unwillingness to settle down (Ong, 2018).

In this study, we aim to find if the assumption of Singaporeans' unwillingness to settle down as the main cause of low fertility rates holds true. In Singapore's society, conservative Asian family values are deeply set in their culture where family cohesion is strongly advocated, and marriage is well-regarded as the strongest level of commitment between a couple (Barr, 2003). Hence, we will use marriage as a measure of willingness to settle down for our research study; is Singapore's annual fertility rate less affected by marriages?

3. Data Cleaning and Transformation

The datasets used in this study had many additional indices and words that had no value to the study. For instance, there was a row which detailed that the dataset was sourced from Singapore Department of Statistics. Since these additional rows were not useful to the study, I removed them by using the drop function and indicated the data frame index numbers. E.g. df = df.drop(df.index [22:26])

There were also other data quality issues in the fertility data set, where some columns had missing values. From earlier time periods like year 1961, there were no data collected and the only data input for no data collected was 'na'. These missing data would be read in the data frames as 'NaN' and disrupt the results of the study. Thus, I should change all ‘na’ value to ‘0’. However, I chose to remove them with function drop.() again as most missing values were before year 1980 and outside the timeline used in this study.

```
df2 = pd.read_csv('birth_fertility_rates.csv', skiprows = 4, index_col = 0)
df2 = df2.drop(df2.index[22 : 26])
#In this dataset, the time series was from 1960 to 2017.
#Since the first dataset of marriages does not have any data from 1960 to 19
#Therefore, we remove the first 20 years in this dataset of fertility rates
df2 = df2.drop(df2.columns[0 : 20], axis = 1)
df2 = df2.drop(df2.columns[-1], axis = 1) #Remove additional unnamed column
```

Code 1:Removing unwanted columns and index

After removing all the unwanted columns and index, we select the variables we want by using loc function and slice the variables by their names. The dataset was also not neatly organised with the years as columns and the variables like fertility rates as rows. To re-organise the dataset in the data frame, I transposed the data frame to invert the columns and rows.

```
df2 = df2.loc[[' Total Fertility Rate (Per Female) ', ' Net Reproduction Rate (Per Female) ']]
#Likewise to the first dataset, I choose the index or rows I want from this dataset for my data
df2 = df2.T
```

Code 2: Slicing and selecting variables before transposing data frame

The variables name from the marriage dataset was also very lengthy and needed to be changed. Using df.rename(), I renamed the variables to a shorter name in the data frame.

```
df = df.rename(index = {' Male General Marriage Rate (Per 1,000 Unmarried Resident Males 15-49) ' : 'Marriage Rates (Males)'})
df = df.rename(index = {' Female General Marriage Rate (Per 1,000 Unmarried Resident Females 15-49) ' : 'Marriage Rates (Females)'})
df = df.T #I transpose the dataframe to make the years e.g. 1985 the index and the Marriage Rates the columns in the dataframe
```

Code 3: Renaming the data frame variables

Variables	index	Total Fertility Rate (Per Female)	Net Reproduction Rate (Per Female)
0	1980	1.82	0.86
1	1981	1.78	0.84
2	1982	1.74	0.82
3	1983	1.61	0.76
4	1984	1.62	0.76
5	1985	1.61	0.76
6	1986	1.43	0.68
7	1987	1.62	0.77
8	1988	1.96	0.93
9	1989	1.75	0.83
10	1990	1.83	0.87
11	1991	1.73	0.83
12	1992	1.72	0.82
13	1993	1.74	0.83
14	1994	1.71	0.82
15	1995	1.67	0.8
16	1996	1.66	0.79
17	1997	1.61	0.77

Output 1: Result of data cleaning fertility dataset

4. Approaches / Methods

We will approach this study with the aim to derive two measures through a combination of linear regression and ratio analysis. The measures are coefficient of correlation and ratio of change in fertility rate to change in marriage rates. With the coefficient of correlation as the linearity and the ratio of change as the magnitude of the relationship, the results from these measures would provide us with sufficient evidence for the link between marriage and fertility rates.

a. Coefficient of correlation

The coefficient of correlation is a statistical measure of the strength and direction of the association between two variables. The range of a coefficient of correlation is between -1.0 to 1.0, depending on the relationship between the variables. A high or low coefficient would mean a direct or inverse relationship and ‘0’ would be no association.

Using the pandas data frame function ‘.corr()’, I could compute the coefficient of correlation between all the column variables in one data frame.

```
#I use the pandas function .co
df_correlation = df3.corr()
#I also calculated the mean of
```

```
print("\nCoefficient of Correlation between Variables from 1980 to 2017:")
print("=" * 150)
print(df_correlation)
```

Code 4: Coefficient of Correlation

In this case, I had calculated the coefficient of correlation between the entire dataset from year 1980 to 2017 and the data frame 'df3' had all the variables from both datasets used. Generating the coefficient of correlation from the pandas function '.corr()', the coefficient of correlation shows that the marriage rates for both male and females has a strong and direct relationship with fertility rates at 0.893408 and 0.923044 respectively.

Even though both male and female marriage rates are positive, it can be noted that the female marriage rates are slightly higher than the male marriage rates. This is an indication that married female Singaporean are more likely to result in higher fertility rates than married Singaporean men.

```
Coefficient of Correlation between Variables from 1980 to 2017:
======================================================================================================================================================
                                 Marriage Rates (Males)  Marriage Rates (Females)  Total Fertility Rate (Per Female)  Net Reproduction Rate (Per Female)
Marriage Rates (Males)                         1.000000                  0.934503                           0.893408                            0.893246
Marriage Rates (Females)                       0.934503                  1.000000                           0.923044                            0.915148
 Total Fertility Rate (Per Female)             0.893408                  0.923044                           1.000000                            0.999171
 Net Reproduction Rate (Per Female)            0.893246                  0.915148                           0.999171                            1.000000
```

Output 2: Coefficient of Correlation

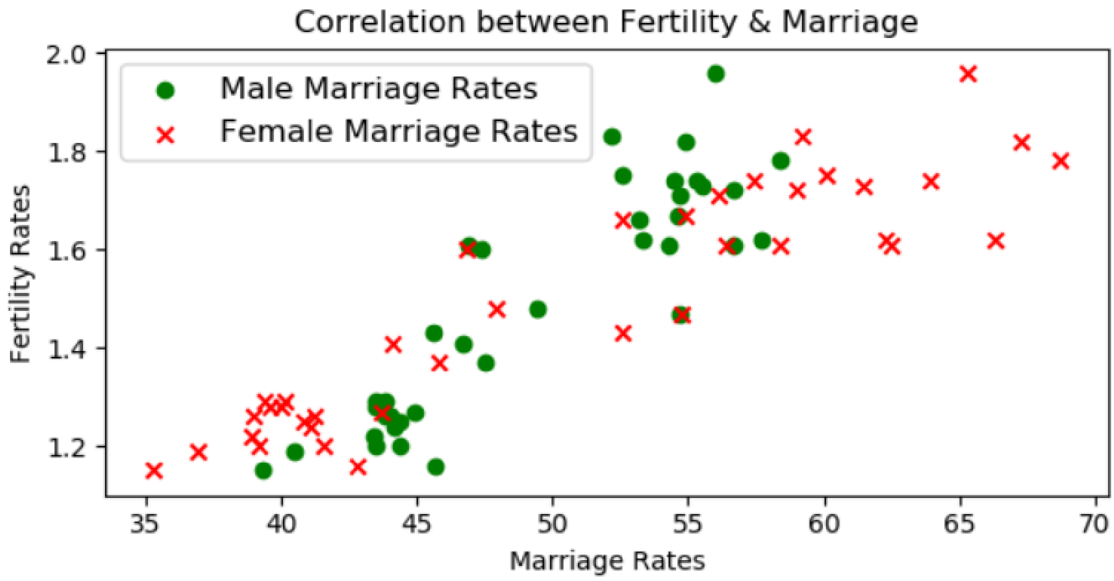


Figure 1: Scatter plot of Marriage vs Fertility Rates

Plotting the marriage rates against the fertility rates on a scatter plot, the positive relationship from the computed coefficient of correlation is further confirmed as a clear uphill, linear pattern is illustrated in the scatter plot. Similar to the coefficient of correlation tabulated, the plots of female marriage rates also look more closely plotted together and not widely spread out, showing a higher correlation with fertility than the male marriage rates.

Overall, we can infer that marriage rates have a clear direct relationship with fertility rates from the coefficient of correlation alone.

b. Ratio of % Change in Fertility Rate to % Change in Marriage Rates

Calculating Changes in Fertility and Marriage Rates

Another measure we could use to evaluate the relationship between the variables is a ratio of change in fertility rate to change in marriage rates to illustrate how sensitive fertility rates are to changes in marriage rates.

We derive this ratio by first finding the changes in fertility and marriage rates over the time period. Using pandas 'diff' function, we find the percentage change in marriage and fertility rates by measuring the difference between the periods. After which, I proceed to add the percentage changes in marriage and fertility rates calculated to another data frame before tabulating the ratios.

Calculating Ratio of Change in Fertility Rate to Change in Marriage Rate

After finding the percentage change in both rates, we use the following formula to find the ratio:

$$\frac{\%\ Change\ in\ Fertility\ Rate}{\%\ Change\ in\ Marriage\ Rate}$$

This ratio is an indicator of the sensitivity of the fertility rate towards the changes in marriage rates and we can infer the degree of the relationship between marriage and fertility. As the yearly changes in both fertility and marriage rates were too minimal, they would not yield very noticeable or substantial results. Therefore, I extended the time interval to 5 years.

```
dfmain['% Change in Marriage Rates (M)'] = dfmain['Marriage Rates (Males)'].astype(float).diff(periods = 1)
dfmain['% Change in Marriage Rates (F)'] = dfmain['Marriage Rates (Females)'].astype(float).diff(periods = 1)
dfmain['% Change in Total Fertility Rate'] = dfmain[' Total Fertility Rate (Per Female) '].astype(float).diff(periods = 1)
dfmain['% Change in Net Reproduction Rate'] = dfmain[' Net Reproduction Rate (Per Female) '].astype(float).diff(periods = 1)

#Using iloc, I proceed to slice the main dataframe dfmain, into two new dataframes df4 andf df5 because they hold different values
#df4 retains the original dataset values and df5 has the % changes in variables over the years
df3.index.name = 'Year'
dfmain.index.name = 'Year'
df4 = dfmain.iloc[:, :4]
df5 = dfmain.iloc[:, 4:]
#I complete the ratio calculation by dividing the calculated % changes in fertility rate over the % changes in marriage rates to derive the ratio
#The computed ratios are added into dataframe df5 as absolute values
df5['Times Change in Fertility for 1% Male Marriage'] = (df5['% Change in Total Fertility Rate']/df5['% Change in Marriage Rates (M)'] *10).abs()
df5['Times Change in Fertility for 1% Female Marriage'] = (df5['% Change in Total Fertility Rate']/df5['% Change in Marriage Rates (F)'] *10).abs()

#Again, I split the dataframe df5 into two new dataframes df6 and df7 to separate my ratios for sensitivity analysis
df6 = df5.iloc[:, :4]
df7 = df5.iloc[:, 4:]
```

Code 5: Calculating ratio of % changes

Based on the ratios generated in output 3 below, we can see how the fertility rates change for every 1% change in male and female marriage rates. For example, if more Singaporean men had married and cause an increase in 1% of male marriage rate, the fertility rate would increase 1.04 times.

Year	Marriage Rates (Males)	Marriage Rates (Females)	Total Fertility Rate (Per Female)	Net Reproduction Rate (Per Female)
1980	54.9	67.3	1.82	0.86
1985	54.3	62.5	1.61	0.76
1990	52.2	59.2	1.83	0.87
1995	54.6	54.9	1.67	0.80
2000	47.4	46.8	1.60	0.76
2005	43.8	39.0	1.26	0.61
2010	39.3	35.3	1.15	0.55
2015	44.2	41.1	1.24	0.60

Sensitivity Analysis

Year	% Change in Marriage Rates (M)	% Change in Marriage Rates (F)	% Change in Total Fertility Rate	% Change in Net Reproduction Rate
1980	NaN	NaN	NaN	NaN
1985	-0.6	-4.8	-0.21	-0.10
1990	-2.1	-3.3	0.22	0.11
1995	2.4	-4.3	-0.16	-0.07
2000	-7.2	-8.1	-0.07	-0.04
2005	-3.6	-7.8	-0.34	-0.15
2010	-4.5	-3.7	-0.11	-0.06
2015	4.9	5.8	0.09	0.05

Year	Times Change in Fertility for 1% Male Marriage	Times Change in Fertility for 1% Female Marriage
1980	NaN	NaN
1985	3.500000	0.437500
1990	1.047619	0.666667
1995	0.666667	0.372093
2000	0.097222	0.086420
2005	0.944444	0.435897
2010	0.244444	0.297297
2015	0.183673	0.155172

Output 3: Calculation of ratios of changes in rates

Plotting the computed ratios in a time series line graph, we can spot a trend of a decreasing ratio over the timeline between 1980 to 2017. Most prominent change is the drastic drop in the sensitivity of fertility rates to male marriages from 3.5 in 1985 to less than 0.5 in 2015. This downward trend of ratios implies that the strong connection between marriage and fertility has weakened over the period of 37 years.

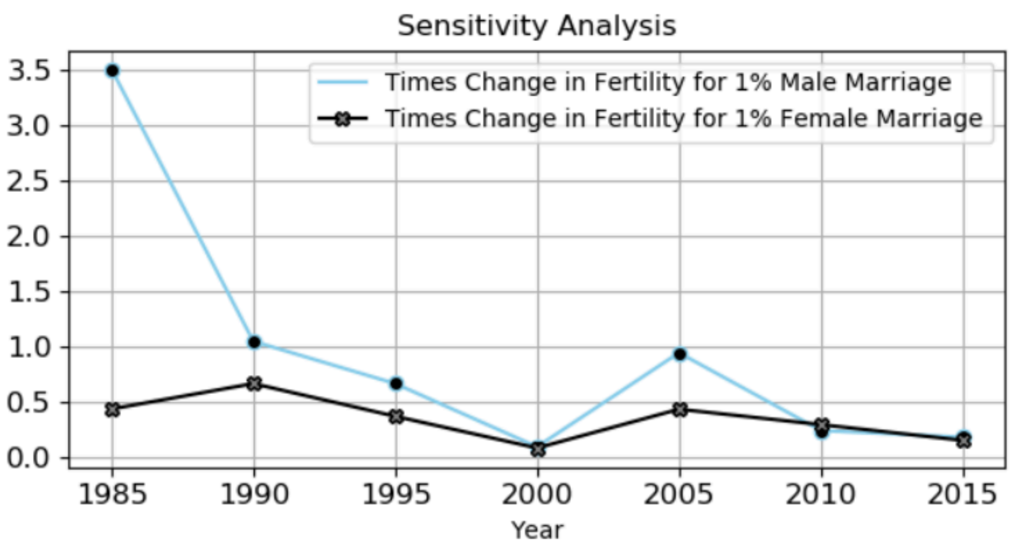


Figure 2: Analysis of Ratios over the years

5. Conclusions

From the measure calculated, we can conclude that the association between marriage and fertility is a direct one. A positive coefficient of correlation has shown that the act of marriage among Singaporeans will result in births and having more Singaporean women getting married would spur this trend even more than married Singaporean men.

On the other hand, the deteriorating trend of sensitivity ratio between the two variables seem to raise an alarmingly trend of weakening effects of marriage on fertility rates. The sensitivity ratios are indicating that marriage may not be the strongest consideration for Singaporeans to have children anymore.

Yet, there are also limitations to this study that needs to be factored into the findings. Limitation includes the assumptions used in calculating the coefficient of correlation. Using Pearson's correlation to compute the coefficient of correlation, we are assuming that marriage and fertility rates in the datasets should always be normally distributed and close to the mean. However, this may not always be the case as it is not entirely impossible to have very low marriage rates in some periods.

The coefficient of correlation also assumes a linear relationship between the two variables. While this may be true when shown in the scatter plot (Figure 1), the sensitivity ratios indicate that the spread of the dots are likely to widen in the future and illustrate a less linear relationship. Hence, the assumption of a linear relationship may not be relevant in future studies.

In summary, this study has shown that marriages in Singapore has led to higher births from 1980 to 2017. However, the influence of marriage in having children is waning and hints that marriage may not equate to children for Singaporeans. Marriage may still encourage Singaporeans to have children, but it is no longer the largest factor in the equation.

Therefore, it is not fair to agree with the Singapore government that Singaporeans' unwillingness to settle down is the main cause of lower fertility rates because marriage is no longer the main consideration for giving birth. It is only one factor in having children and there are possibly other factors in the consideration for giving births, which the government could investigate further.

REFERENCES:

Ong, J. (2018). Singapore's fertility rate at new 7-year low of 1.16: Josephine Teo. Retrieved from https://www.channelnewsasia.com/news/singapore/singapore-total-fertility-rate-new-low-1-16-10002558

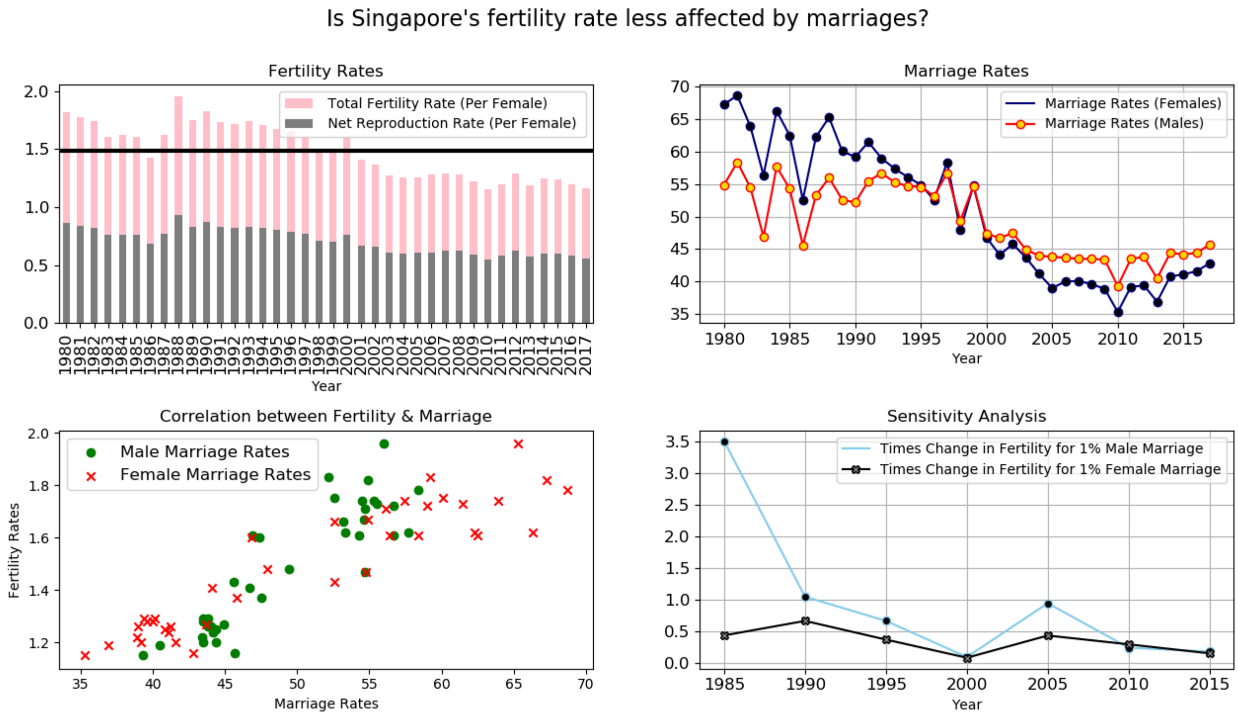
Sin, Y. (2018). Govt aid alone not enough to raise birth rate: Minister. Retrieved from https://www.straitstimes.com/singapore/govt-aid-alone-not-enough-to-raise-birth-rate-minister

Yeung, W., & Hu, S. (2018). Marriage, families under stress as norms change. Retrieved from https://www.straitstimes.com/opinion/marriage-families-under-stress-as-norms-change

Kotecki, P. (2018). 10 countries at risk of becoming demographic time bombs. Retrieved from https://www.businessinsider.sg/10-countries-at-risk-of-becoming-demographic-time-bombs-2018-8/?r=US&IR=T

Barr, M. D. (2003). *Cultural Politics and Asian Values*. London: Routledge.

APPENDIX A



Is Singapore's fertility rate less affected by marriages?
Fertility Rates
Total Fertility Rate (Per Female)
Net Reproduction Rate (Per Female)
Year
Marriage Rates
Marriage Rates (Females)
Marriage Rates (Males)
Year
Correlation between Fertility & Marriage
Male Marriage Rates
Female Marriage Rates
Fertility Rates
Marriage Rates
Sensitivity Analysis
Times Change in Fertility for 1% Male Marriage
Times Change in Fertility for 1% Female Marriage
Year